

Journal club

Model for predicting *EGFR* mutation status in lung cancer

Cite as: Nguyen Ho L, Le TV. Model for predicting EGFR mutation status in lung cancer. *Breathe* 2019; 15: 340–342.

Commentary on:

Wang S, Shi J, Ye Z, *et al.* Predicting *EGFR* mutation status in lung adenocarcinoma on CT image using deep learning. *Eur Respir J* 2019; 53: 1800986

Context

Lung cancer is a leading cause of cancer-related mortality worldwide, with an estimated 1.38 million deaths annually [1]. The approach to diagnosis and treatment has changed considerably, with developments such as 1) screening of lung cancer to identify the early stage lesions, 2) various sampling methods to diagnose the histopathological features of the lung tumour, and 3) changing from conventional chemotherapy to molecular targeted therapy. In the era of precision medicine, targeted therapy consistent with specific oncogenic mutation, such as tyrosine kinase inhibitor (TKI) treatment in lung adenocarcinoma with epidermal growth factor receptor (EGFR) mutation, has an important role.

The most common form of lung cancer is nonsmall cell lung cancer (NSCLC), of which two major pathologic groups are adenocarcinoma and squamous cell carcinoma. The EGFR mutation appears mainly in lung adenocarcinoma, which is predominant in Asian women without

history of smoking [1]. TKI treatment for lung adenocarcinoma with sensitising EGFR mutation improves progression-free survival and quality of life compared to conventional chemotherapy [2].

However, EGFR mutation testing for lung adenocarcinoma has specific barriers. It is difficult to obtain a specimen to analyse the mutation, because the majority of lung cancer patients present with advanced stage and are unsuitable for invasive sampling procedures. EGFR testing can show false-negative results in some cases [1]. To resolve these issues, plasma testing for EGFR mutation or using a predictive model for EGFR mutation status could be useful in making a clinical decision. WANG *et al.* [3] showed the potential application of a deep learning model in predicting the EGFR mutation status in lung adenocarcinoma.

Methods

Inclusion criteria were: 1) histologically confirmed primary lung adenocarcinoma; 2) pathologic examination of tumour specimens already carried out with proven records of EGFR mutation status; 3) pre-operative contrast-enhanced computed tomography (CT) data obtained. Exclusion criteria were: 1) clinical data including age, gender and stage was missing; 2) pre-operative treatment was received; 3) the duration between CT examination and subsequent surgery exceeded 1 month. This study included 844 patients, 603 of whom were



CrossMark



© ERS 2019

 @ERSpublications

A predictive model using available chest CT images could better assess the presence of *EGFR* mutations. This model may also identify biopsy false-negative results. <http://bit.ly/2MR0mpe>

from Shanghai Pulmonary Hospital (the primary cohort used to develop the deep learning model) and 241 of whom were from Tianjin Medical University (the validation cohort). The EGFR mutation of specimens from surgical tumour resection was determined by an amplification refractory mutation system with a human EGFR mutations detection kit (Beijing ACCB Biotech Ltd, Beijing, China).

Development of the deep learning model was done through model training, in which the first 20 convolutional layers were trained by 1.28 million natural images from the ImageNet dataset and the last four convolutional layers by 14926 CT images from lung adenocarcinoma tumours in the primary cohort. In applying this deep learning model, the cubic region of interest containing the whole tumour was identified without image segmentation, resized to 64×64 pixels by third-order spline interpolation in each CT slice, and fed into the deep learning model. The probability of the tumour being EGFR-mutant was given directly.

The study used SPSS Statistics 21 (IBM, Armonk, NY, USA). The independent samples *t*-test for the mean age and the deep learning score between two groups with EGFR-mutant and EGFR-wild type, the Chi-squared test for categorical variables such as gender and tumour stage, and the DeLong test for the receiver operating characteristic (ROC) curves between various models were used to determine the difference.

Main results

This deep learning model showed good predictive performance in both the primary cohort and the validation cohort, with area under ROC curve (AUC) 0.85 (95% CI 0.83–0.88) and 0.81 (95% CI 0.79–0.83), respectively. The deep learning score revealed a significant difference between two groups with EGFR-mutant and EGFR-wild type regardless of tumour stage. The decision curve analysis showed the usefulness of this predictive model in clinical practice.

This study also compared the deep learning model with others using clinical characteristics, semantic features, or quantitative “radiomic” features, and concluded that the deep learning model obtained better predictive performance in both the primary cohort and the validation cohort.

With the deep learning visualisation method, this study would support clinicians in identifying suspicious tumour area, in particular, in relation to EGFR mutation, to lower the false-negative rate of EGFR testing.

Commentary

WANG *et al.* [3] revealed an important feature of using artificial intelligence to predict EGFR mutation. Gene mutation alters the biological process, resulting in

clinical manifestation. Therefore, characteristics of the malignancy condition could be useful in predicting the presence of an oncogene [4]. Clinical manifestation, histopathological features and lung tumour shape were used to develop the predictive model of EGFR mutation in lung adenocarcinoma patients [2–7]. The lung tumour possessed complex features called “radiomic” features (shape, intensity, texture and wavelet) which were difficult for radiologists to identify directly [5, 7]. The application of artificial intelligence, particularly deep learning, has potential value in analysing the complex features of CT images of lung tumours [8]. In the article, WANG *et al.* [3] compared the deep learning model to others to find whether it would provide better prediction. However, whether a model combining the clinical manifestation, pathologic feature, and radiologic feature prediction could outperform the predictive performance of this deep learning model requires further investigation.

Moreover, it is unknown whether NSCLC patients could be treated with TKIs with a positive outcome based on this deep learning model rather than EGFR mutation testing in real-world clinical practice. By 2005, the empirical use of erlotinib for refractory NSCLC had been accepted widely in Europe, the USA and elsewhere, but afterwards, the randomised IUNO trial showed no improvement of overall survival [9]. The National Comprehensive Cancer Network 2017 guidelines stopped the recommendation of using erlotinib as switch maintenance treatment for EGFR-wild type NSCLC. Applying this predictive model, there is evidence to use EGFR TKIs when EGFR testing is not directly accessible. However, this deep learning model only predicts the EGFR mutation status and does not provide more information relating the type of EGFR mutation (sensitising or resistance to TKI) which is of great importance in clinical practice [10].

This deep learning model indicated the suspicious tumour area which could be most relevant to the investigation of EGFR mutation [3]. Low quality biopsy specimens or low-sensitivity EGFR testing [1] might lead to a false-negative EGFR result in positive-predicted patients. Resampling at the suspicious tumour area should be considered.

The major source of histopathology in these predictive models was from surgical tumour resection [2, 3, 7]. However, a high proportion of patients with lung cancer were ineligible for surgical therapy and the real-life diagnosis was conducted mainly through small biopsy procedures (bronchoscopy, transthoracic lung biopsy). KIM *et al.* [4] reported the association between the clinical pathological features with EGFR mutation *via* small biopsies. An important problem of small biopsy specimens is whether they are representative of the entire tumour with regards to EGFR mutation status. Published studies show conflicting results about the intratumour genetic/molecular heterogeneity

of EGFR mutation [11–13]. With the concordance rate of 68–97.1% between biopsy tissue and the whole tumour for EGFR mutation status [11], we could modestly accept the representative role of biopsy tissue. However, this uncertainty justifies the need to study the predictive model of EGFR mutation in lung cancer patients with small biopsy in the future.

We believe that the extensive application of this deep learning model needs to be questioned because the study population was only Chinese. Genetic mutations in lung cancer exhibit racial and ethnic differences [14]. For example, studies have shown that the prevalence of EGFR mutation was higher in Asian than non-Asian patients. The potential differentiation of “radiomic” characteristics on chest CT images of lung tumours might not be transferable to other cohorts from other parts of the world. It is important that local resources of CT images

are readily available to train this model again and re-evaluate it before using.

Implications for practice

The deep learning model of WANG *et al.* [3] predicts EGFR mutation status using available chest CT images, particularly in patients from whom it is not possible to obtain tumour tissue samples. This can support clinicians in making treatment decisions.

This model also suggests a solution for the false-negative result of EGFR testing. Tumour biopsy at the suspicious tumour areas that are more susceptible to the EGFR mutation should be considered.

It is essential to conduct the predictive model study in lung cancer patients diagnosed by small biopsy because of the high prevalence of unresectable lung cancer patients.

Affiliations

Lam Nguyen Ho, Thuong Vu Le

Dept of Internal Medicine, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam.

Acknowledgements

We thank Dr. Le Hong Van for assistance in editing English of this manuscript.

Conflict of interest

None declared.

References

1. Ellison G, Zhu G, Moulis A, *et al.* EGFR mutation testing in lung cancer: a review of available methods and their use for analysis of tumour tissue and cytology samples. *J Clin Pathol* 2013; 66: 79–89.
2. Liu Y, Kim J, Qu F, *et al.* CT features associated with epidermal growth factor receptor mutation status in patients with lung adenocarcinoma. *Radiology* 2016; 280: 271–280.
3. Wang S, Shi J, Ye Z, *et al.* Predicting EGFR mutation status in lung adenocarcinoma on CT image using deep learning. *Eur Respir J* 2019; 53: 1800986.
4. Kim HJ, Choi EY, Jin HJ, *et al.* Relationship between EGFR mutations and clinicopathological features of lung adenocarcinomas diagnosed via small biopsies. *Anticancer Res* 2014; 34: 3189–3195.
5. Cao Y, Xu H. A new predictive scoring system based on clinical data and computed tomography features for diagnosing EGFR-mutated lung adenocarcinoma. *Curr Oncol* 2018; 25: e132–e138.
6. Velazquez ER, Parmar C, Liu Y, *et al.* Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res* 2017; 77: 3922–3930.
7. Zhang L, Chen B, Liu X. Quantitative biomarkers for prediction of epidermal growth factor receptor mutation in non-small cell lung cancer. *Transl Oncol* 2018; 11: 94–101.
8. Zhao W, Yang J, Ni B, *et al.* Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. *Cancer Med* 2019; 8: 3532–3543.
9. Itchins M, Clarke S, Pavlakis N. Do EGFR tyrosine kinase inhibitors (TKIs) still have a role in EGFR wild-type pre-treated advanced non-small cell lung cancer (NSCLC)?—the shifting paradigm of therapeutics. *Transl Lung Cancer Res* 2018; 7: Suppl. 1, S39–S45.
10. Gazdar AF. Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. *Oncogene* 2009; 28: Suppl. 1, S24–S31.
11. de-Sousa VML, Carvalho L. Heterogeneity in lung cancer. *Pathobiology* 2018; 85: 96–107.
12. Yatabe Y, Matsuo K, Mitsudomi T. Heterogeneous distribution of EGFR mutations is extremely rare in lung adenocarcinoma. *J Clin Oncol* 2011; 29: 2972–2977.
13. Taniguchi K, Okami J, Kodama K, *et al.* Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. *Cancer Sci* 2008; 99: 929–935.
14. Schabath MB, Cress D, Munoz-Antonia T. Racial and ethnic differences in the epidemiology and genomics of lung cancer. *Cancer Control* 2016; 23: 338–346.